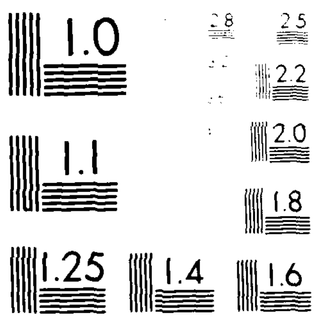
AD-A116 083 STANFORD UNIV CA INFORMATION SYSTEMS LAB F/G 12/1
TIME-VARIANT AND TIME-INVARIANT LATTICE FILTERS FOR NONSTATIONA--ETC(U)
MAY 82 T KAILATH F49620-79-C-0058
UNCLASSIFIED AFOSR-TR-82-0481 NL

1 OF 1
AD A
116083

END
DATE
FILMED
07-82
DTIC


1.0
2.8
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

AFOSR-TR- 82-0481

(11)

AD A116083

TIME-VARIANT AND TIME-INVARIANT LATTICE FILTERS FOR NONSTATIONARY PROCESSES[†]

T. Kailath

Information Systems Laboratory
Department of Electrical Engineering
Stanford University
Stanford, CA 94305

ABSTRACT

The structure of second-order processes is exposed by specification of whitening filters and modeling filters, or equivalently by Cholesky decompositions of the covariance matrix and its inverse. We shall show that these filters can be obtained as a cascade of lattice sections, each specified by a single so-called reflection coefficient parameter. For stationary processes, the reflection coefficient will be time-invariant. For nonstationary processes we can use the displacement rank concept either to find a simple time-update formula for the reflection coefficients or to replace them by a time-invariant vector reflection coefficient of size governed by the displacement rank of processes.

These results are obtained in a quite direct way by using a geometric (Hilbert-space) formulation of the problem, combined with old results of Yule (1907) on update formulas for partial correlation coefficients and of Schur (1917) and Szego (1939) on the classical moment problem.

MAY 1982

[†] This work was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command under Contract AF49-620-79-C-0058, by the U.S. Army Research Office, under Contract DAAG29-79-C-0215 and the Joint Services Program under Contract DAAG29-81-K-0057.

1. INTRODUCTION

In recent years there has been considerable interest in lattice filters for signal processing. Though such structures had been well studied in network theory as, for example, in the wave digital filters of Fettweis (see, e.g., Fettweis et al. (1974)) and especially in the cascade synthesis of multiport networks (Dewilde (1969)), the first applications in on-line adaptive signal processing were apparently made by Itakura and Saito (1971) in the field of speech analysis. Lattice filter models were also familiar in geophysical signal processing as 'layered earth models' (see, e.g., Robinson (1967)); Burg ((1970), unpublished) explicitly suggested lattice filter models for the implementation of a spectral estimation technique based on the maximum _ntropy principle (Burg (1967), see also Alam (1978)). The lattice filters used in these works were based on the assumption of an underlying *stationary* process, and though in fact the filters were applied to deterministic and nonstationary processes, it was believed that then the lattice solutions were suboptimal (see e.g., Makhoul (1977)). Morf and Vieira (see Morf, Vieira and Lee (1977), Vieira (1977)) were the first to show that the optimal solutions to certain (so-called prewindowed) nonstationary model fitting problems could be obtained in lattice form, though with time-variant lattice parameters (reflection coefficients). These lattice filters demonstrated excellent tracking capability and rapid convergence (see, e.g., Morf and Lee (1978) (1979), Satorius and Pack (1981), Hodgkiss and Presley (1981)). A recent thesis by Lee (1980) and a paper by Lee, Morf and Friedlander (1981) present a number of efficient normalized versions of the prewindowed lattice (which they call ladder) algorithms. A forthcoming paper by Porat, Friedlander and Morf (1982) presents a similar comprehensive analysis of the so-called "covariance" lattice algorithms.

In Section 3 of this paper we shall show that these results on the prewindowed and covariance algorithms can be embedded into a more general theory of time-variant lattice filters for nonstationary stochastic processes. We shall start in Section 2 by showing that lattice filters can be set up for determining the innovations of any (second-order) process; however, the computational effort is $O(N^3)$, where N is the number of observed random variables, and this is the same as for any other method (e.g., direct Gram-Schmidt orthogonalization) of determining the innovations. Of course, the lattice structure has various advantages: modularity, local interconnections, simple checks for stability, reduced sensitivity to parameter and roundoff errors. We may note that some of these nice properties arise from the fact that the lattice algorithms are a form of Modified Gram-Schmidt algorithm, which is known to have better numerical properties than the direct Gram-Schmidt procedure (see, e.g., Stewart (1973)). However, our major point is that order of magnitude reductions in the computational burden can be obtained by introducing the concept of displacement rank as a measure of nonstationarity; in particular, we shall use the displacement rank to obtain a general time-update formula for the reflection coefficients. The derivation is quite direct and elegant, calling upon an old identity, due to Yule (1907), on partial correlation coefficients and a geometric characterization, due to Delosme and Morf (1980), of displacement rank via so-called "pinning vectors", which were introduced by Sidhu and Kailath (1975) in connection with the Chandrasekhar equations (Kailath (1973), Kailath, Sidhu,Morf (1973)). While inputs from a number of people (especially D. T. L. Lee, M. Morf, D. R. Morgan, J-M. Delosme) contributed to the development, our presentation follows one put together by H. Lev-Ari, and to appear in his thesis (Lev-Ari, (1982)).

In the final section of this paper, we shall show how the displacement rank can be exploited to obtain a different kind of simplification—a time-invariant lattice filter but with somewhat more complicated sections, having $\alpha-1$ delays per section rather than just one, where α is the displacement rank. This result was first obtained as a consequence of the generalized Levinson algorithm derived by Friedlander et al. (1978), (1979). The original derivation was quite lengthy and relied heavily on insights from the state-space Chandrasekhar equations; in the meantime, through contributions from Dewilde, Vieira, Porat, Genin, Delosme, Dym and others, and especially Lev-Ari, a much clearer view has emerged. Certain classical results of Schur on tests for the positivity of moment matrices, when combined with the displacement rank concept, turn out to provide a simple, yet general and insightful, view of the topic. We shall present an outline of this approach in Section 4, while a detailed exposition and further results will appear in the thesis of H. Lev-Ari (1982).

2. LATTICE RECURSIONS FOR GENERAL NONSTATIONARY PROCESSES

We shall be studying an indexed (by time) collection of "vectors" $\{y_i,\ i=0,1,\ldots,t,\ldots\}$ in some Hilbert space, with given inner products

$$\langle y_i, y_j \rangle = R_{i,j} \ . \tag{1}$$

and such that the Gramian matrix

$$\mathbf{R} = [R_{i,j}] \tag{2}$$

is a symmetric positive definite matrix. We generally have in mind that the $\{y_i\}$ are random variables in a probability space where the inner product is defined by expectation,

$$\langle y_i, y_j \rangle = E\{y_i y_j^*\} \ .$$

where the asterisk denotes complex conjugation (or Hermitian transpose, when applied to matrix quantities). However, this stochastic interpretation is not at all necessary.

We shall denote

$$||y_i||^2 := \langle y_i, y_i \rangle$$

and shall use bars to denote normalized quantities, e.g.,

$$\bar{y}_i = ||y_i||^{-1} y_i \ . \tag{3}$$

The structure of the collection of vectors $\{y_i\}$ will be explored by studying the family of finite-order residuals

$$e_{m,t} := y_t - \hat{y}_{t|U} \ , \tag{4}$$

where

$$U = \{y_{t-m}, \ldots, y_{t-1}\} \tag{5}$$

and

$$\hat{y}_{t|U} = \text{the projection of } y_t \text{ on the space spanned by the set } U . \tag{6}$$

We shall call

$$e_{m,t} = \text{the } m\!-\!th \text{ order forward residual at time } t . \tag{7}$$

We note that the residuals

$$e_{t,t} , \ t = 0,1,2,...$$

will be the so-called *innovations process* associated with the $\{y_t, t=0,1,...\}$ which is usually the chief object of attention in least-squares estimation theory; to obtain lattice filters, however, it is convenient to imbed these desired quantities in the larger family of residuals $\{e_{m,t}, m \le t, t=0,1,...\}$.

The structure of this family can be exposed by first seeking, for each fixed t, how to determine order-updates of these residuals, i.e., knowing $e_{m,t}$ we shall try to determine

$$e_{m+1,t} = y_t - \hat{y}_{t|\{U, y_{t-m-1}\}}$$

in some convenient way. It is reasonable to seek to use our knowledge of $e_{m,t}$ by making the orthogonal decomposition

$$\{U, y_{t-m-1}\} = U \oplus \{r_{m,t-1}\}$$

where

$$r_{m,t-1} := y_{t-m-1} - \hat{y}_{t-m-1|U} , \tag{8}$$

the $m\!-\!th$ *order backward residual at* $t-m-1$.

Then we can write

$$\begin{aligned} \hat{y}_{t|\{U, y_{t-m-1}\}} &= \hat{y}_{t|U} + \hat{y}_{t|r_{m,t-1}} \\ &= \hat{y}_{t|U} + \langle y_t, \bar{r}_{m,t-1} \rangle \bar{r}_{m,t-1} . \end{aligned} \tag{9}$$

It follows that we can write

$$
\begin{aligned}
e_{m+1,t} &= e_{m,t} - <y_t, \bar{r}_{m,t-1}> \bar{r}_{m,t-1} \\
&= e_{m,t} - <e_{m,t}, \bar{r}_{m,t-1}> \bar{r}_{m,t-1} \\
&= ||e_{m,t}||(\bar{e}_{m,t} - <\bar{e}_{m,t}, \bar{r}_{m,t-1}> \bar{r}_{m,t-1}) \\
&= ||e_{m,t}||(\bar{e}_{m,t} - k_{m+1,t}\bar{r}_{m,t-1})
\end{aligned}
$$

where we used the fact that $r_{m,t-1} \perp U$ to obtain the second equality, and where we defined

$$
k_{m+1,t} := <\bar{e}_{m,t}, \bar{r}_{m,t-1}> \quad . \tag{10}
$$

For reasons given later, such quantities will be called "reflection coefficients".

To also normalize $e_{m+1,t}$, we need to compute its norm, for which we note that the last equality yields

$$
||e_{m+1,t}||^2 = ||e_{m,t}||(I - k_{m+1,t}k^*_{m+1,t})||e_{m,t}||^*
$$

so that

$$
||e_{m+1,t}|| = ||e_{m,t}||(I - k_{m+1,t}k^*_{m+1,t})^{1/2} \tag{11}
$$

Then we can rewrite the order-update formula for the forward residuals in normalized form as

$$
\bar{e}_{m+1,t} = (I - k_{m+1,t}k^*_{m+1,t})^{-1/2}(\bar{e}_{m,t} - k_{m+1,t}\bar{r}_{m,t-1}) \tag{12}
$$

This, of course, leaves us with the problem of getting $\bar{r}_{m,t-1}$. But a similar recursion can be set up for it. Thus and more briefly, we can write

$$
\begin{aligned}
r_{m+1,t} &= y_{t-m-1} - \hat{y}_{t-m-1|\{U, y_t\}} \\
&= r_{m,t-1} - <y_{t-m-1}, \bar{e}_{m,t}> \bar{e}_{m,t} \\
&= ||r_{m,t-1}||(\bar{r}_{m,t-1} - <\bar{r}_{m,t-1}, \bar{e}_{m,t}> \bar{e}_{m,t}) \\
&= ||r_{m,t-1}||(\bar{r}_{m,t-1} - k^*_{m+1,t}\bar{e}_{m,t})
\end{aligned}
$$

This yields

$$
||r_{m+1,t}|| = ||r_{m,t-1}||(I - k^*_{m+1,t}k_{m+1,t})^{1/2}
$$

so that

$$\bar{r}_{m+1,t} = (I - k^{*}_{m+1,t}k_{m+1,t})^{-1/2}(\bar{r}_{m,t-1} - k_{m+1,t}\bar{e}_{m,t}) \tag{13}$$

The recursions (12) and (13) can be pictorially represented as in Figure 1(a) as a typical "lattice" section. Moreover, we see that we can put the sections together as in Figure 1(b), to get a cascaded structure, with each lattice section specified by a reflection coefficient,

$$k_{m+1,t} = \langle \bar{e}_{m,t}, \bar{r}_{m,t-1} \rangle, \quad m = 0,1,\dots .$$

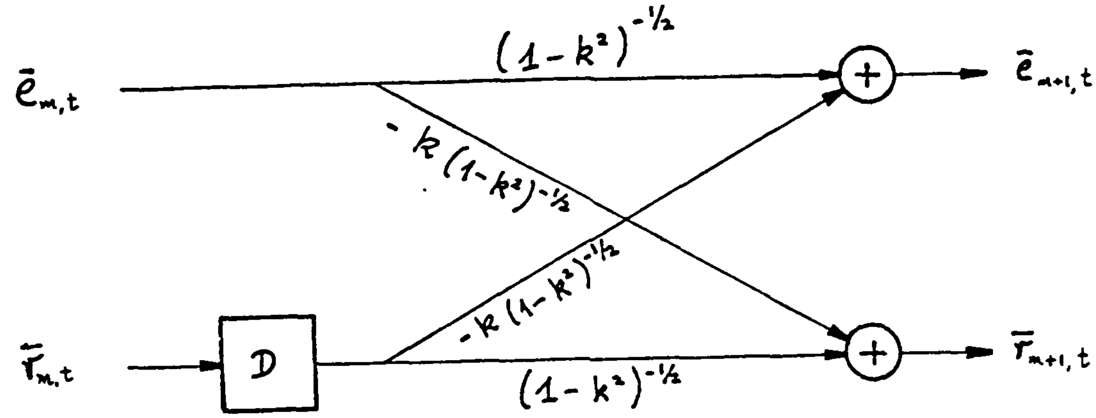


Figure 1(a). Lattice Section

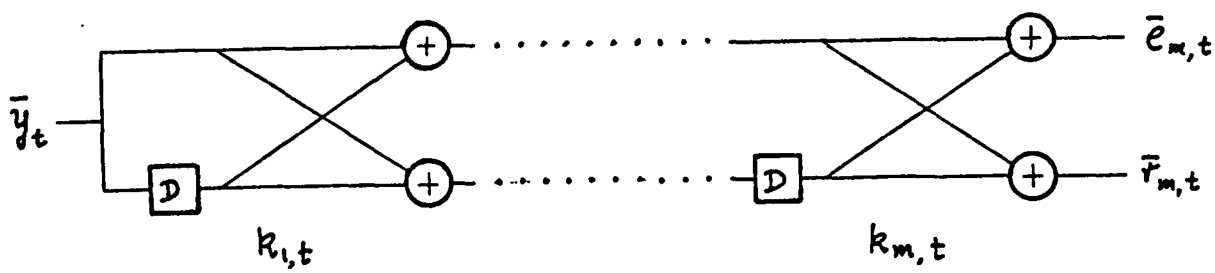


Figure 1(b). Lattice Filter with Time-Varying Sections

Note that the inputs to the first section are

$$\bar{e}_{0,t} = \bar{y}_t = ||y_t||^{-1} y_t = \bar{r}_{0,t} \tag{14}$$

and that the first section has reflection coefficient

$$k_{1,t} = \langle \bar{e}_{0,t}, \bar{r}_{0,t-1} \rangle = \langle \bar{y}_t, \bar{y}_{t-1} \rangle = R_{t,t}^{-1/2} R_{t,t-1} R_{t-1,t-1}^{-1/2} \tag{15}$$

The name reflection coefficient is partly justified by the fact that, by Schwarz's inequality,

$$||k_{m+1,t}|| \leq ||\bar{e}_{m,t}|| \, ||\bar{r}_{m,t-1}||^{*} = 1 \ . \tag{16}$$

The Stationary Case

A more detailed justification is obtained from closer study of the "stationary" case, where stationarity means that the inner products are invariant under shift in the indices, i.e.,

$$\langle y_i, y_j \rangle = \langle y_{i-1}, y_{j-1} \rangle , \quad i = 1,2,\ldots \tag{17}$$

In this case, it follows that the reflection coefficients $\{k_{m+1,t}\}$ are *independent* of t, so that we have a cascade lattice filter with constant or *time-invariant* sections. Physical realizations of such time-invariant filters (and in particular of their inverses) have been studied in some detail in geophysics (see, e.g., Robinson (1967), Claerbout (1976)), and in speech analysis (see, e.g., Wakita (1973), Markel and Gray (1976)), and provide the real justification for the name reflection coefficients. We shall not pursue this further here, but may refer to Paper No.   by Benveniste earlier in this volume.

Simplifications in the Nonstationary Case

In the nonstationary case, when (17) does not hold, it seems unavoidable that the reflection coefficients must be time-variant. However, it is reasonable that the complexity of the time-variation should depend upon the degree of nonstationarity, measured in some sense. It turns out that the concept of displacement rank provides a meaningful way of classifying nonstationary processes, in that for a process with displacement rank α, each reflection coefficient can be updated with $O(\alpha)$ multiplications. For N observations and N reflection coefficients, this requires $O(N^2\alpha)$ multiplications as compared to the $O(N^3)$ that would be required if we just used the general formulas given above without attention to the displacement rank.

The time-update formulas can be compactly stated:

$$k_{m+1,t} = (1-\eta_{m,t}\eta_{m,t}^{*})^{1/2}k_{m+1,t-1}(1-\mu_{m,t-1}\mu_{m,t-1}^{*})^{1/2} + \eta_{m,t}\mu_{m,t-1}^{*} \qquad (18)$$

where $\{\eta,\mu\}$ are α-dimensional row vectors obeying the recursions (Yule's Parcor identity)

$$\eta_{m+1,t} = F\{\eta_{m,t}, k_{m+1,t}, \mu_{m,t-1}\} \qquad (19)$$
$$\mu_{m+1,t} = F\{\mu_{m,t-1}, k_{m+1,t}^{*}, \eta_{m,t}\} \qquad (20)$$

where the function $F(\,)$ is defined as

$$F\{A,B,C\} = (1-BB^{*})^{-1/2}(A-BC^{*})(1-CC^{*})^{-*/2} \qquad (21)$$

In fact, we may note that the first equation is a rearrangement of the formula.

$$\begin{aligned} k_{m+1,t-1} &= F\{k_{m+1,t}, \eta_{m,t}, \mu_{m,t-1}\} \qquad (22)\\ &= (I-\eta\eta^{*})^{-1/2}(k_{m+1,t}-\eta\mu^{*})(I-\mu\mu^{*})^{-*/2} \end{aligned}$$

where the subscripts for μ and η have been omitted for simplicity.

The displacement rank itself is given as

$$\alpha = \text{rank of } \{\mathbf{R} - Z\mathbf{R}Z^{*}\} \qquad (23)$$

where $\mathbf{R}$ is the covariance or Gramian matrix (2) and Z is the lower shift operator with ones on the first subdiagonal and zeros elsewhere.

The three recursions (18)-(20) can be pictorially represented as in Figure 2. The block named D denotes a "delay operator" and the numbers indicated denote implementation of the respective equations.

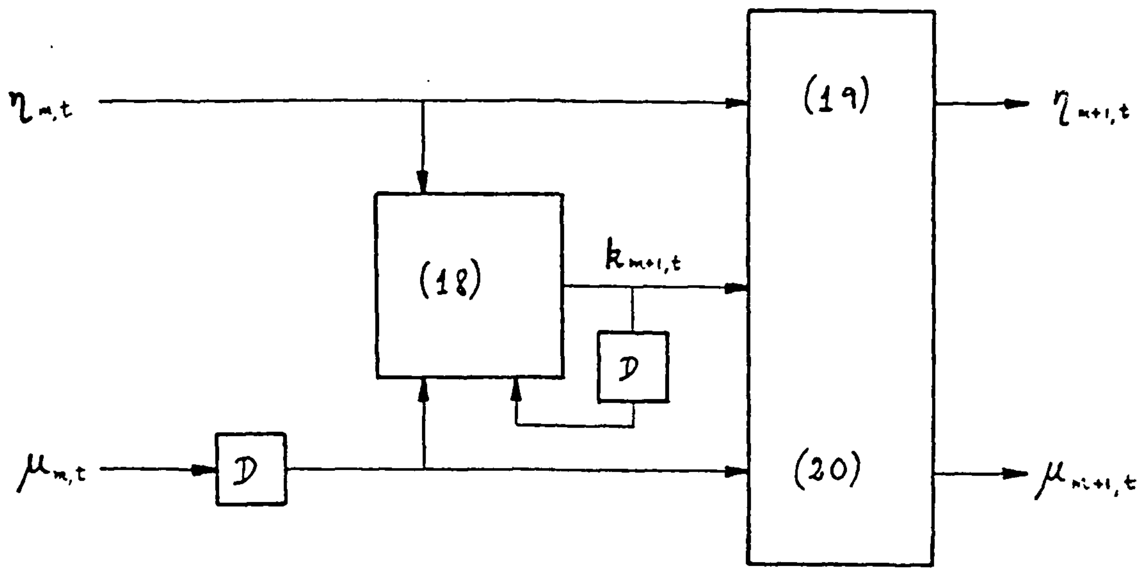


Figure 2. Representation of the Time-Update Calculations

Time-Invariant Implementations

Perhaps surprisingly, the displacement rank can be used to reduce the complexity in a different way--by allowing completely time-invariant gains but of a higher dimension. That is, we shall still have a cascade of lattice sections, but each section will be defined by an $(\alpha-1)$-dimensional row vector rather than a scalar--see Figure 3. These row vectors will be called generalized Schur coefficients.

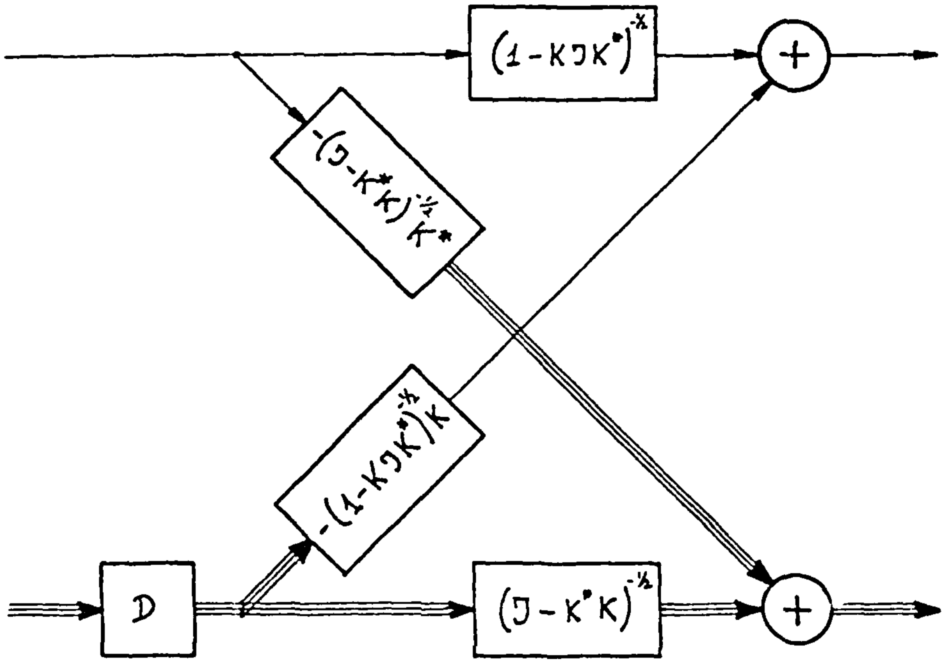


Figure 3. Generalized Time-Invariant Lattice Section

As shown, the input to the 'Delay' is a $(\alpha-1)\times 1$ signal; the 'Delay' itself is $z^{-1}I_{\alpha-1}$, where $I_{\alpha-1}$ denotes an identity matrix of dimension $(\alpha-1)\times(\alpha-1)$. It can be replaced by $\left[\frac{z-\alpha_i}{1-\alpha_i^* z}\right]^{-1}\times I_{\alpha-1}$, to yield an ARMA lattice structure. The generalized Schur coefficients $\{K_i\}$ are row vectors of dimension $\alpha-1$, satisfying

$$1 - K_i J K_i^* \geq 0 \tag{24}$$

where α is the displacement rank of the covariance matrix $\mathbf{R}$ of and J is defined through the signature of $\mathbf{R}$ as:

$$\operatorname{sgn}\{\mathbf{R} - Z\,\mathbf{R}\,Z^*\} = \{1, -J\} \tag{25}$$

Thus J is an $(\alpha-1)\times(\alpha-1)$ diagonal matrix with entries $+1$ or -1. Relation (24) can be compactly written by using the J-norm notation:

$$||K_i||_J \leq 1$$

To conclude this introduction, we should also remark that we shall show that the generalized Schur coefficients $\{K_i\}$ are not unique: different set of $\{K_i\}$ can be associated with a given covariance function, even in the stationary case!

Also, a given set of $\{K_i\}$ can correspond to several different covariance functions, all "congruent" to each other (see Section 4).

3. DERIVATION OF THE TIME-UPDATE FORMULAS

The time update formulas (18)-(20) will follow very easily by combining certain results on partial correlation coefficients and on displacement ranks. These will be presented in the first two subsections. The third subsection will contain the proofs of (18)-(29) and the final subsection will give an application to the so-called exact deterministic least-squares algorithm for "pre-windowed" data sequences.

3.1 Parcor Coefficients and Yule's Identity

The reflection coefficients as defined by (10) can be identified as partial correlation (Parcor) coefficients. Before proceeding with their analysis, it will be helpful to introduce some notation and derive some general properties of such coefficients.

Let a,b be vectors in some Hilbert space. We shall denote the *correlation coefficient* of a and b as

$$\rho(a,b) = <\bar{a},\bar{b}> = ||a||^{-1} \cdot <a,b> \cdot ||b||^{-1} \tag{26}$$

where $\bar{a}=||a||^{-1}a$ is the normalized version of a. Given any set V of vectors, let $\tilde{a}_{|V}$, or more simply a_V, denote the *residual* error in estimating a from V. Thus

$$a_V := a - <a,\bar{V}>\bar{V} \tag{27}$$

where $\bar{V}$ is any orthonormal basis for the linear space spanned by the set V. Then, we shall define the *partial correlation coefficient of* $\{a,b\}$ *given* V, a concept first introduced by Yule (1907), as

$$\begin{aligned} \rho_V(a,b) &= \rho(\bar{a}_V,\bar{b}_V) \\ &= ||a_V||^{-1}<a_V,b_V>||b_V||^{-*} \end{aligned} \tag{28}$$

In this notation, the reflection coefficient $k_{m+1,t}$ defined in (10) can be written as,

$$k_{m+1,t} = \rho_U(y_t, y_{t-m-1}) \tag{29}$$

where we recall that (cf. (5))

$$U = \{y_{t-m}, \ldots, y_{t-1}\}$$

We can write

$$\begin{aligned} a_V &= a - \langle a, \bar{V} \rangle \bar{V} \\ &= ||a|| \cdot (\bar{a} - \langle \bar{a}, \bar{V} \rangle \bar{V}) \\ &= ||a|| \cdot (\bar{a} - \rho(a, V)\bar{V}) \end{aligned}$$

and check that

$$||a_V|| = ||a|| \cdot (I - \rho(a,V)\rho^*(a,V))^{1/2}$$

so that

$$\begin{aligned} \bar{a}_V &= ||a_V||^{-1} \cdot a_V \\ &= (I - \rho(a,V)\rho^*(a,V))^{-1/2}(\bar{a} - \rho(a,V)\bar{V}) \ . \end{aligned}$$

With a similar expression for $\bar{b}_V$, we can then deduce that

$$\begin{aligned} \rho_V(a,b) &= \rho(a_V, b_V) = \langle \bar{a}_V, \bar{b}_V \rangle \\ &= (I - \rho(a,V)\rho^*(a,V))^{-1/2} \langle (\bar{a} - \rho(a,V)\bar{V}), (\bar{b} - \rho(b,V)\bar{V}) \rangle \\ &\cdot (I - \rho(b,V)\rho^*(b,V))^{-1/2} \\ &= (I - \rho(a,V)\rho^*(a,V))^{-1/2}(\rho(a,b) - \rho(a,V)\rho^*(b,V))(I - \rho(b,V)\rho^*(b,V))^{-*/2} \end{aligned} \tag{30}$$

or more compactly

$$\begin{aligned} \rho_V(a,b) &= F\{A, B, C\} \\ &= (I - BB^*)^{-1/2}(A - BC^*)(I - CC^*)^{-*/2} \end{aligned} \tag{31}$$

where

$$A = \rho(a,b) \ ; \ B = \rho(a,V) \ ; \ C = \rho(b,V) \ .$$

Now we shall seek a formula for modifying a partial correlation coefficient

$\rho_U(a,b)$ when the set U is changed to $\{U,V\}$

For this, note first that the space spanned by $\{U,V\}$ can be written as a direct sum $=\{U \oplus V_U\}$, where V_U is orthogonal to U. Now $a_{\{U,V\}}$ can be computed as follows:

First estimate a from U, with error a_U; then estimate a_U from V_U, with final overall error of $(a_U)_{V_U}$. Therefore we have a nice formula, which the reader should thoroughly assimilate:

$$a_{\{U,V\}} = (a_U)_{V_U}$$

Then we can write,

$$\begin{aligned}\rho_{\{U,V\}}(a,b) &= \rho(a_{\{U,V\}}, b_{\{U,V\}}) \\ &= \rho((a_U)_{V_U}, (b_U)_{V_U}) \\ &= \rho_{V_U}(a_U, b_U)\end{aligned}$$

which by using (31), we can write as

$$\begin{aligned}\rho_{\{U,V\}}(a,b) &= \rho_{V_U}(a_U, b_U) \\ &= F\{\rho(a_U, b_U), \rho(a_U, V_U), \rho(b_U, V_U)\} \\ &= F\{\rho_U(a,b), \rho_U(a,V), \rho_U(b,V)\} \qquad (32)\end{aligned}$$

This basic result first appeared (in different notation) in Yule's original paper (1907) and therefore we shall call it Yule's Parcor Identity.

3.2 Displacement Ranks, Differential Generators, and Pinning Aggregates

We turn now to the other key ingredient of the time-update formula. Let us denote

$$\mathbf{R}_{0:N} = \left[R_{i,j}\right]_{i,j=0}^{N} \text{ where } R_{i,j} = \langle y_i, y_j \rangle \qquad (33)$$

If we define a block column vector

$$\mathbf{Y}_{0:N} = [y_0, y_1, \ldots, y_N]^* \quad . \tag{34}$$

then we can write

$$\mathbf{R}_{0:N} = \langle \mathbf{Y}_{0:N}, \mathbf{Y}_{0:N} \rangle \quad . \tag{35}$$

Now we introduce the displacement operator

$$\lrcorner\, \mathbf{R}_{0:N} = \mathbf{R}_{0:N} - Z\, \mathbf{R}_{0:N} Z^* \tag{36a}$$

where

$$Z = \begin{bmatrix} 0 & & & & \\ 1 & \cdot & & 0 & \\ & \cdot & \cdot & & \\ & 0 & \cdot & \cdot & \\ & & & 1 & 0 \end{bmatrix} \tag{36b}$$

The displacement rank is defined as

$$\alpha = \text{rank}(\lrcorner\, \mathbf{R}) \tag{37}$$

Since $\mathbf{R}_{0:N}$ is Hermitian, so will be $\lrcorner\,\mathbf{R}_{0:N}$. Therefore, it has real eigenvalues, say q_+ strictly positive, and q_- strictly negative, so that

$$\alpha = q_+ + q_- \tag{38}$$

Then we can write $\lrcorner\,\mathbf{R}_{0:N}$, although nonuniquely, as

$$\mathbf{R}_{0:N} - Z\, \mathbf{R}_{0:N}\, Z^* = G_{0:N}\, \Sigma\, G_{0:N}^* \tag{39}$$

where

$$\Sigma = \begin{bmatrix} I_{q+} & 0 \\ & \\ 0 & -I_{q-} \end{bmatrix} \tag{40}$$

and

$$G_{0:N} = \begin{bmatrix} \cdot & \cdot & g_0 & \cdot & \cdot \\ \cdot & \cdot & g_1 & \cdot & \cdot \\ & & \vdots & & \\ \cdot & \cdot & g_N & \cdot & \cdot \end{bmatrix} \tag{41}$$

$(G_{0:N}, \Sigma)$ will be called a 'Differential Generator' of $\mathbf{R}_{0:N}$, since it is not hard to see that $\mathbf{R}_{0:N}$ can be uniquely recovered from knowledge of $\{G_{0:N}, \Sigma\}$.

An interesting geometric interpretation can be associated with (39), rewritten as

$$\mathbf{R}_{0:N} - G_{0:N} \Sigma G_{0:N}^{*} = Z \mathbf{R} Z^{*} \tag{42}$$

Delosme and Morf (1980) showed that we could find a collection, X, of vectors, α in number, in an extended Hilbert space, such that

$$<X,X> = \Sigma \text{ and } G_{0:N} = <\mathbf{Y}_{0:N}, X> \tag{43}$$

We may remark that such so called "pinning aggregates" were first introduced by Sidhu and Kailath (1975) to obtain a geometric derivation of the state-space Chandrasekhar equations.

Now, in terms of X, we can express (42) in the suggestive form,

$$<\mathbf{Y}_{0:N}, \mathbf{Y}_{0:N}> - <\mathbf{Y}_{0:N}, X><X,X>^{-1}<\mathbf{Y}_{0:N}, X>^{*} = <Z \mathbf{Y}_{0:N}, Z \mathbf{Y}_{0:N}> \tag{44}$$

For, noticing that the residual

$$(\mathbf{Y}_{0:N})_X = \mathbf{Y}_{0:N} - <\mathbf{Y}_{0:N}, X><X,X>^{-1}X$$

equation (44) shows that

$$<(\mathbf{Y}_{0:N})_X, (\mathbf{Y}_{0:N})_X> = <Z \mathbf{Y}_{0:N}, Z \mathbf{Y}_{0:N}> \tag{45}$$

This equality allows us to set up an isometry (i.e., a norm-preserving isomorphism) between the spaces

$$\{\text{space of vectors } (Y_{0:N})_X\} \sim \{\text{space of } Z \mathbf{Y}_{0:N}\} \tag{46}$$

In particular, we have

$$(y_0)_X \sim 0 \text{ and } (y_t)_X \sim y_{t-1} \tag{47}$$

Now returning to our problem, we recall that

$$
\begin{aligned}
k_{m+1,t} &= \langle \bar{e}_{m,t}, \bar{r}_{m,t-1} \rangle \\
&= \rho[(y_t)_{\mathcal{Y}}, (y_{t-m-1})_{\mathcal{Y}}] \\
&= \rho_{\mathcal{Y}}(y_t, y_{t-m-1})
\end{aligned}
$$

where

$$U = \{y_{t-m}, \ldots, y_{t-1}\}$$

It will be convenient to introduce an operator D such that

$$Dy_t = y_{t-1}, \quad t \geq 1 \ . \tag{48a}$$

$$DU = \{y_{t-m-1}, \ldots, y_{t-1}\} \tag{48b}$$

With this notation, what we are seeking in the time update formula is a relation between

$$k_{m+1,t} = \rho_{\mathcal{Y}}(y_t, y_{-m-1}) \text{ and } k_{m+1,t-1} = \rho_{D\mathcal{Y}}(Dy_t, Dy_{t-m-1}) \tag{49}$$

3.3 Derivation of Time-Update Formulas

With the help of the isometry and Yule's Parcor identity, we can derive the time-update formulas as follows.

Using first the isometry (48) and then Yule's identity (32), we can write

$$
\begin{aligned}
k_{m+1,t-1} &= \rho_{D\mathcal{Y}}(Dy_t, Dy_{t-m-1}) \\
&= \rho_{\mathcal{Y}_X}((y_t)_X, (y_{t-m-1})_X) \\
&= \rho_{\{\mathcal{Y},X\}}(y_t, y_{t-m-1}) \\
&= F\{\rho_{\mathcal{Y}}(y_t, y_{t-m-1}), \rho_{\mathcal{Y}}(y_t, X), \rho_{\mathcal{Y}}(y_{t-m-1}, X)\} \\
&= F\{k_{m+1,t}, \eta_{m,t}, \mu_{m,,t-1}\}
\end{aligned}
\tag{49}
$$

where we have defined

$$\eta_{m,t} = \rho_{\mathcal{Y}}(y_t, x) \text{ and } \mu_{m,t-1} = \rho_{\mathcal{Y}}(y_{t-m-1}, x) \tag{50}$$

We can readily obtain recursions for these quantities as well. Thus

$$
\begin{aligned}
\eta_{m+1,t} &= \rho_{\{y_{t-m-1}, \mathcal{Y}\}}(y_t, x) \\
&= F\{\rho_{\mathcal{Y}}(y_t, x), \rho_{\mathcal{Y}}(y_t, y_{t-m-1}), \rho_{\mathcal{Y}}(x, y_{t-m-1})\} \\
&= F\{\eta_{m,t} k_{m+1,t}, \mu^{\bullet}_{m,t-1}\}
\end{aligned}
\tag{51}
$$

and finally,

$$\mu_{m+1,t} = \rho_{\{y,y_t\}}(y_{t-m-1},x)$$
$$= F\{\rho_{y}(y_{t-m-1},x),\rho_{y}(y_{t-m-1},y_t),\rho_{y}(x,y_t)\}$$
$$= F\{\mu_{m,t-1},k^*_{m+1,t},\eta^*_{m,t}\} \tag{52}$$

Equations (49), rewritten to express $k_{m+1,t}$ in terms of $k_{m+1,t-1}$, combined with equations (51)-(52) are just the time-update formulas presented earlier in Section II. Note how immediately they follow from the isometry and Yule's identity.

3.4 Application to the Time-Series Pre-Windowed Case

In this section, we shall apply the theory developed so far to a deterministic least-squares problem, studied by Morf, Vieira and Lee (1977) and Lee, Morf and Friedlander (1981).

(i) Statement of the Problem

Given a series of *samples*

$$y(0),y(1),\ldots,y(t),\ldots,y(N)$$

in the so-called pre-windowed method, we try to fit an *nth* order autoregression to it by choosing coefficients $\{1,a_1,\ldots,a_n\}$ to minimize the following quantity:

$$\sum_{t=0}^{N} e_n^2(t) \tag{53}$$

where

$$e_n(t) = y(t) + a_1 y(t-1) + \cdots + a_n y(t-n) \tag{54a}$$

with

$$y(-k) \equiv 0\,,\quad k \geq 1\ . \tag{54b}$$

The assumption (54b) means that we are assuming that samples prior to

$t=0$ are zero, hence the name "pre-windowed". The residuals $e_n(t)$ can be written in a matrix form as below:

$$[a_n \quad \cdots \quad a_1 \quad 1]\begin{bmatrix} 0 & 0 & & y(0) & y(1) & . & y(N-n) \\ . & . & & . & & & . \\ \vdots & 0 & \cdot & & & & \vdots \\ 0 & y(0) & \cdot & & & & . \\ y(0) & y(1) & \cdot & . & . & . & y(N) \end{bmatrix}$$

$$= [e_n(0) \quad \cdots \quad e_n(N)] \tag{55a}$$

or more compactly

$$\mathbf{A}_n \cdot \mathbf{Y}(n) = \mathbf{e}_{n,N} \tag{55b}$$

The solution to our minimization problem can then be seen to be the solution of the *normal equations*.

$$\mathbf{A}_n \mathbf{R}_n = [0 \quad \cdots \quad R_n^e] \tag{56}$$

where

$$\mathbf{R}_n = \mathbf{Y}(n)\mathbf{Y}^*(n) \quad \text{and} \quad R_n^e = \min \sum_{t=0}^{N} e_n^2(t) = ||\mathbf{e}_{n,N}||^2 .$$

Note that $\mathbf{R}_n$ will not in general be a Toeplitz matrix (unless, in fact, $y(N-n+1) = 0 = \cdots = y(N)$).

(ii) Solution of the Problem

We can cast this deterministic least-squares problem into our previous Hilbert space framework as follows:

Define our Hilbert space elements as row vectors

$$y_t = [0 \quad \cdots \quad 0 \ y(0) \quad \cdots \quad y(t)] \quad \text{for } 0 \le t \le N \tag{57}$$

of some fixed length T, $T > N$. And take as inner products,

$$\langle y_t, y_s \rangle = y_t \cdot y_s^*$$

$$= [0 \quad \cdots \quad 0 \ y(0) \quad \cdots \quad y(t)][0 \quad \cdots \quad 0 \ y(0) \quad \cdots \quad y(s)]^*$$

$$= \sum_{i=0}^{t\wedge s} y(t-i)y^*(s-i) \tag{58}$$

where $t \wedge s$ stands for the smaller of t and s. Then, we see from (55a) that choosing $\{1, a_1, \ldots, a_n\}$ to minimize $\sum_{t=0}^{N} e_n^2(t)$ is equivalent to finding the best prediction of the Hilbert space vector y_N given n earlier vectors $\{y_{N-1}, \ldots, y_{N-n}\}$.

We can now use our earlier general result as follows: First, we must find the displacement rank α and the pinning vector X. By our definition of inner product, we have

$$\mathbf{R}_{o:N} = \begin{bmatrix} y_o \\ y_1 \\ \cdot \\ \cdot \\ y_N \end{bmatrix} [y_o^*, \ \ldots, \ y_N^*]$$

Thus,

$$\begin{aligned} \underline{\ \ }|\mathbf{R}_{o:N} &= \mathbf{R}_{o:N} - Z\mathbf{R}_{o:N}Z^* \\ &= [y_t y_s^* - y_{t-1} y_{s-1}^*]_{t,s=0:N}\ , \quad y_{-i} \equiv 0 \ \text{ for } \ i \geq 1 \end{aligned} \tag{59}$$

Note that the expression in the square brackets denotes the (t,s)-th element of $\underline{\ \ }|\mathbf{R}_{o:N}$, with t and s taking values from 0 to N as indicated in the subscript. Substituting $y_t := [0 \ \cdots \ 0 \ y(0) \ \cdots \ y(t)]$, we get

$$\begin{aligned} \underline{\ \ }|\mathbf{R}_{o:N} &= [y(t)y^*(s) - y(t-T)y^*(s-T)]_{t,s=o:N} \\ &= [y(t)y^*(s)]_{t,s=o:N} \quad \text{since} \quad y(-i) = 0 \quad \text{for} \quad i \geq 1\ . \end{aligned} \tag{60}$$

Consider the following $n \times n$ submatrix of $\underline{\ \ }|\mathrm{R}_{o:N}$, denoted briefly as $[\underline{\ \ }|\mathbf{R}]_{N-n:N}$:

$$\begin{aligned} [\underline{\ \ }|\mathbf{R}]_{N-n:N} &= [y(t)y^*(s)]_{t,s=N-n:N} \\ &= \begin{bmatrix} y(N-n) \\ \vdots \\ y(N) \end{bmatrix} [1][y^*(N-n) \ \cdots \ y^*(N)] \end{aligned} \tag{61}$$

$$= G_{N-n:N} \cdot \Sigma \cdot G^*_{N-n:N} \ , \quad \text{say,}$$

We see that the displacement rank, which equals the dimension of Σ, is

$$\alpha = 1 \ . \tag{62}$$

Also that the pinning vector is

$$X = [0 \quad \cdots \quad 0 \quad 1] \ , \tag{63}$$

because

$$<[y^*_{N-n}, \quad \ldots, \quad y^*_N]^*, X> \; = \; [y^*(N-n), \quad \ldots, \quad y^*(N)]^*$$

$$= G_{N-n:N}$$

With $\{\alpha, X\}$, we can, as before, build a lattice filter with n sections, where the $(m+1)$-th section has reflection coefficient

$$k_{m+1,t} = \rho_U(y_t, y_{t-m-1}) \ , \qquad U = \{y_{t-m}, \ \ldots, \ y_{t-1}\} \ .$$

To update this quantity, we need $\{\eta, \mu\}$ where

$$\eta_{m,t} = \rho_U(y_t, X) \quad \text{and} \quad \mu_{m,t-1} = \rho_U(y_{t-m-1}, X) \ .$$

In the present problem

$$<y_t, X> \; = \; y(t) \quad \text{and} \quad <y_{t-m-1}, X> \; = \; y(t-m-1) \ . \tag{64}$$

Then some calculation will show that

$$\eta_{m,t} = \bar{e}_m(t) \cdot c \ , \quad \mu_{m,t} = \bar{r}_m(t-1) \cdot c \tag{65}$$

where the proportionality factor c arises from the absence of a scaling factor (corresponding to a so called second normalization in Lee et al. (1981)). The point really is that the lattice sections need only to propagate the sample values $\{\bar{e}_m(t), \ \bar{r}_m(t-1)\}$ and not the row vectors $\{\bar{\mathbf{e}}_{m,t}, \ \bar{\mathbf{r}}_{m,t-1}\}$. We refer to the paper of Lee et al. (1981) for further details on the prewindowed lattice form.

4. GENERALIZED SCHUR ALGORITHMS AND TIME INVARIANT LATTICE FILTERS

We shall begin by considering tests for positivity (positive-definiteness) of a given matrix, starting with Toeplitz matrices, which can arise as covariance matrices of stationary processes. A simple test is shown using Schur coefficients (= reflection coefficients) obtained via the standard Schur algorithm. Then a generalized Schur algorithm will be developed yielding generalized Schur coefficients which will give positivity tests for non-Toeplitz matrices. Then, the interpretation of the Schur algorithm is given, using concepts of spectral functions and Schur invariance. Later, we shall consider these results in the context of nonstationary process, and discuss prediction filters and the generalized Szegö-Levinson algorithm. We shall briefly also mention orthogonal polynomial and Christoffel-Darboux formulas.

4.1 Positivity of Toeplitz Matrices

Consider a stationary stochastic process $\{y_t\}$ with an associated Toeplitz covariance matrix

$$\mathbf{R}_{0:N} \begin{bmatrix} R_0 & R_1 & R_2 & \cdot & \cdot & \cdot & R_N \\ R_1 & \cdot & \cdot & & & & \\ R_2 & \cdot & \cdot & \cdot & & & \\ \cdot & & \cdot & \cdot & \cdot & & \\ \cdot & & & \cdot & \cdot & \cdot & \\ \cdot & & & & \cdot & \cdot & R_1 \\ R_N & & & & & R_1 & R_0 \end{bmatrix} , \quad R_0 = 1$$

There are many criteria for testing whether a given matrix is positive definite, based on computing eigenvalues or by reduction to a sum of squares, etc., but we are seeking more special tests that exploit the special (Toeplitz) structure of the matrix. Such tests are in fact available, based on certain function-theoretical characterizations first developed in the early part of this century.

Let us temporarily assume that we have an infinite sequence

$$R_0, R_1, \ldots$$

and define a function (assuming $R_0 = 1$)

$$C(z) = 1 + 2\sum_{1}^{\infty} R_i z^i \quad . \tag{66}$$

Then we have the following characterization.

Lemma (Caratheodory (1910)): The matrices $\{R_{0:N}, N = 0,1,\ldots\}$ will be nonnegative definite if and only if $C(z)$ is analytic and has nonnegative real part in the unit disc.

Testing if $\mathrm{Re}\ C(z) \geq 0$ can be replaced by the generally simpler problem of testing for boundedness by making the transformation

$$S(z) = [1 - C(z)][1 + C(z)]^{-1} \tag{67}$$

Then it is not hard to see that $C(z)$ is analytic and $\mathrm{Re}C(z) \geq 0$ in $|z| \leq 1$ if and only if

$$S(z) \text{ is analytic and } |S(z)| \leq 1 \text{ in } |z| \leq 1 . \tag{68}$$

Such a function will be called a *Schur function.*

Lemma: (Schur (1917)): Let $S_0(z) = S(z)$ and define

$$S_{i+1}(z) = \frac{1}{z} \frac{S_i(z) - S_i(0)}{1 - S_i(0)S_i(z)}, \quad i = 0,1,\ldots \tag{69}$$

Then $S(z)$ is a Schur function if and only if $\{S_i(z)\}$ are also Schur functions. Furthermore, let us define what we shall call Schur coefficients

$$k_{i+1} = S_i(0) \quad . \tag{70}$$

Then $S(z)$ is a Schur function if and only if $|k_i| \leq 1$ for $1 \leq i \leq \infty$.

These $\{k_i\}$ are in fact the reflection coefficients used to construct the time-invariant filter for a stationary process (cf. Section 2), a connection first noted by Dewilde, Vieira and Kailath (1978).

Some examination will show that the computation of $\{k_i, 1 \le i \le N\}$, needs knowledge only of $\{R_0 = 1, R_1, \ldots, R_N\}$ and that therefore we in fact have a test for any finite matrix $\mathbf{R}_{0:N}$:

$$\mathbf{R}_{0:N} \ge 0 \text{ if and only if } |k_i| \le 1, \; 1 \le i \le N . \qquad (71)$$

Moreover, further reflection will show the reasonableness of the following result and corollary.

Theorem: The covariance matrix $\mathbf{R}_{0:N}$ can be extended by adding terms $\{R_i, i > N\}$ corresponding to added reflection coefficients k_i, such that $|k_i| \le 1, i > N$.

Corollary: The 'maximum entropy" extension is achieved by the choice $\{k_i = 0, i > N\}$

Before leaving this section, we remark that the reflection coefficients $\{k_i, 1 \le i \le N\}$ can be computed by operations on the first N coefficients of the functions $\{1 + C(z), 1 - C(z)\}$, in particular on the matrix

$$G_{0:N} = \begin{bmatrix} 1 & 0 \\ R_1 & -R_1 \\ R_2 & -R_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ R_N & -R_N \end{bmatrix}$$

The details of the algorithm will be shown later (Section 4.4).

4.2 Positivity of Non-Toeplitz Matrices

Suppose now that the process is not stationary so that,

$$\langle y_t, y_s \rangle \ne \langle y_{t-1}, y_{s-1} \rangle \; .$$

Then the associated covariance matrix will not be Toeplitz and we might ask how we could extend the previous tests for the positive definiteness of $\mathbf{R}_{0:N}$. We could try to introduce Caratheodory and Schur functions of 2-variables. For example,

$$C(z,\omega) = R_{0,0} + 2\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} R_{i,j} z^i \omega^j \tag{72a}$$

$$S(z,\omega) = \frac{R_{0,0} - C(z,\omega)}{R_{0,0} + C(z,\omega)} \tag{72b}$$

But, nothing satisfactory seems to have been found in this way, at least so far

However, we shall show that the concept of displacement ranks and differential generators, as introduced in Section 3.2, does allow us to make progress.

4.3 Differential Generators of R :

We recall from Section 3.2 that we can then write, for nonunique $G_{0:N}$,

$$\underline{|}\, \mathbf{R}_{0:N} = \mathbf{R}_{0:N} - Z\, \mathbf{R}_{0:N} Z^* = G_{0:N}\, \Sigma\, G_{0:N}^* \tag{73}$$

where $G_{0:N}$ is an $(N+1)\times\alpha$ matrix with rows $\{g_0, g_1, \ldots, g_N\}$, say, and Σ is a signature matrix for $\underline{|}\,\mathbf{R}_{0:N}$,

$$\Sigma = \begin{bmatrix} I_{q+} & 0 \\ 0 & -I_{q-} \end{bmatrix}$$

$G_{0:N}$ is nonunique, because any product $G_{0:N}U$ with a Σ-unitary matrix U will also satisfy (73). One consequence is that we can always choose, for convenience,

$$g_0 = [R_{0,0}^{1/2} \;\; 0 \;\; \cdots \;\; 0]\,. \tag{74}$$

This choice is a natural one because the reader can check, it allows us to make the first column of $G_{0:N}$ identical to the first column of $\mathbf{R}_{0:N}$. If for some reason we have a generator without this property, an appropriate Σ-unitary matrix U can also be chosen to ensure it--we shall give the precise construction a little later (see Sec. 4.4, Step 1b).

Having $G_{0:N}$, we can now associate a complex function of a *single* variable with a non-Toeplitz matrix $\mathbf{R}$,

$$G(z) = \sum_{i=0}^{N} g_i z^i + O(z^{N+1}) \tag{75}$$

Note that this function is unlike the extended Caratheodory/Schur functions, which are functions of *two* variables. To gain some insight into this function $G(z)$, let us examine it for the stationary case.

$G(z)$ in The Toeplitz Case

When $\mathbf{R}_{0:N}$ is Toeplitz, we see that

$$\nabla \mathbf{R}_{0:N} = \begin{bmatrix} R_0 & R_1 & \cdots & R_N \\ R_1 & R_0 & \cdot & \cdot \\ \vdots & & \ddots & \vdots \\ \vdots & & \ddots & R_1 \\ R_N & \cdots & R_1 & R_0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & R_0 & R_1 \cdots & R_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & R_1 \\ 0 & R_{N-1} & \cdots & R_0 \end{bmatrix}$$

$$= \begin{bmatrix} R_0 & R_1 & \cdots & R_N \\ R_1 & & & \\ \vdots & & O & \\ R_N & & & \end{bmatrix} \tag{76}$$

Therefore $\alpha=2$, $q_+=q_-=1$ and

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

We can also check that

$$G_{0:N} = \frac{1}{R_{0,0}^{1/2}} \cdot \begin{bmatrix} R_0 & 0 \\ R_1 & -R_1 \\ R_2 & -R_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ R_N & -R_N \end{bmatrix} \tag{77}$$

satisfies $R_{0:N} = G_{0:N} \Sigma G_{0:N}^T$. Assume $R_0 = 1$. Then, recalling the definition of the Caratheodory function,

$$C(z) = 1 + 2\sum_{i=1}^{N} R_i z^i + O(z^{N+1})$$

we see that

$$G(z) = \begin{bmatrix} \frac{1 + C(z)}{2} & \frac{1 - C(z)}{2} \end{bmatrix} .$$

Note that the entries of $G(z)$ are precisely those defining the Schur function $S(z)$ in the Toeplitz case, which was introduced in a rather mysterious way then. Here a clue begins to appear to establish a positivity-test for the non-Toeplitz covariance matrix. For reasons of space we shall forego examination of how the Schur algorithm in the stationary case can be recast in terms of $G(z)$; instead we shall directly present the (very easy to describe) general Schur algorithm:

4.4 General Schur Algorithm

In the non-Toeplitz case, $G(z)$ will not have any such striking form as in (53), but it will nevertheless turnout to yield generalized Schur coefficients in a nice way. Also the procedure includes the special method for the Toeplitz case. Note first that in the general case, $G_{0:N}$ will have α-columns, which we partition as follows:

$$G_{0:N} = \left[\begin{array}{c|cccc} \sigma_0 & 0 & 0 & \cdots & 0 \\ \gamma_1 & & \Gamma_1 & & \\ \cdot & & \cdot & & \\ \cdot & & \cdot & & \\ \cdot & & \cdot & & \\ \gamma_N & & \Gamma_N & & \end{array}\right] \tag{78a}$$

Here we have exploited the nonuniqueness of $G_{0,N}$ to make the first row $[\sigma_0\ 0\ \cdots\ 0]$, where

$$R_{0,0} = \sigma_0^2 = g_0 \Sigma g_0^* > 0 \qquad (78b)$$

In a corresponding way, we partition Σ as follows:

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & -J \end{bmatrix} \qquad (79a)$$

where

$$-J = \begin{bmatrix} I_{q_+-1} & 0 \\ 0 & -I_{q_-} \end{bmatrix} . \qquad (79b)$$

Thus Σ and J are $\alpha \times \alpha$ and $(\alpha-1)\times(\alpha-1)$ matrices. In order to present the Schur algorithm, we need a further definition.

Σ-Unitary Matrices: Matrices of the following type play an important role in the ensuing theory:

$$\Theta(K,J) = \begin{bmatrix} (1 - KJK^*)^{-1/2} & 0 \\ 0 & (J - K^*K)^{-1/2} \end{bmatrix} \begin{bmatrix} 1 & -KJ \\ -K^* & I_{\alpha-1} \end{bmatrix} \qquad (80)$$

We can verify that

$$\Theta \Sigma \Theta^* = \Sigma \qquad (81)$$

thus being called Σ-unitary matrices. In fact, it can be shown that an arbitrary Σ-unitary matrix will have the form

$$\begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \Theta(K,J) \text{ where } U_1 \text{ is unitary and } U_2 \text{ is } J\text{-unitary} . \qquad (82)$$

The Generalized Schur Algorithm:

Start with $G_{0:N}$ in the 'proper' form, i.e., such that the first row g_0 of $G_{0:N}$ is of the form

$$g_0 = [\sigma_0 \ 0 \ \cdots \ 0], \quad \sigma_0 > 0 .$$

Step Ia. Shifting: Form a matrix $\tilde{G}^{(1)}$ with one less row than G by shifting up the rows Γ_i.

$$\tilde{G}^{(1)} = \begin{bmatrix} \sigma_0 & \Gamma_1 \\ \gamma_1 & \Gamma_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \gamma_{N-1} & \Gamma_N \end{bmatrix} \tag{83}$$

On

$$G(z) = \sum_{i=0}^{\infty} g_i z^i ,$$

this operation amounts to forming

$$\tilde{G}^{(1)}(z) = G(z) \begin{bmatrix} I & 0 \\ 0 & zI_{\alpha-1} \end{bmatrix}^{-1} \tag{84}$$

Step Ib. Renormalization: Find $\Theta(K_1, J)$ so that $\tilde{G}^{(1)}\Theta^{-1}(K_1, J)$ is again proper, i.e.,

$$G^{(1)} = \tilde{G}^{(1)}\Theta^{-1}(K_1, J) = \left[\begin{array}{c|ccc} \sigma_1 & 0 & \cdots & 0 \\ \gamma_1^{(1)} & & \Gamma_1^{(1)} & \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ \gamma_{N-1}^{(1)} & & \Gamma_{N-1}^{(1)} & \end{array}\right] \tag{85}$$

we shall show that this can be achieved by choosing

$$K_1 = -\sigma_0^{-1}\Gamma_1 J . \tag{86}$$

Proof of (86): From (85) we need $\Theta(K_1, J)$ to be such that

$$\begin{aligned} \text{the 1st row of } \tilde{G}^{(1)} &= [\sigma_1 \ 0 \ \cdots \ 0]\Theta(K_1, J) \\ &= \sigma_1 \text{ 1st row of } \Theta(K_1, J) \\ &= \sigma_1(1 - K_1 J K_1^*)^{-1/2}[1 \quad -K_1 J] . \end{aligned}$$

Comparing with (83), we get

$$\sigma_0 = \sigma_1(1 - K_1JK_1^*)^{-1/2} \text{ and } \Gamma_1 = -\sigma_0K_1J \text{ . Hence } \ldots$$

Steps II, III, etc: The two steps as before, namely, shifting and renormalization, are repeated until $\tilde{G}^{(m)}$ is a matrix with all columns zero except, perhaps, for the first column. At this stage, no more renormalization is required and all future $\{K_i\}$ will be zero. Note that $m \leq N$, the size of the original matrix and if $m=N$, $\tilde{G}^{(N)}$ will consist of only a single row.

In this way , we can associate a set of $1\times(\alpha-1)$ vector coefficients $\{K_1, K_2, \ldots, K_N\}$ with $G_{0:N}$ and hence with $\mathbf{R}_{0:N}$. The following result holds.

Theorem: The matrix $\mathbf{R}_{0:N}$ is positive-definite if and only if the set of Schur coefficients associated with it satisfy

$$1 - K_iJK_i^* > 0 \quad \text{for } i = 1, \ldots, N \ . \tag{87}$$

In the Toeplitz case $\alpha=2$, $J=1$ so that $\{K_i\}$ have to be scalars of magnitude less than one, as we stated earlier. We have a striking generalization here.

We shall not give a formal proof of this theorem here, but we shall outline the underlying ideas in the next section.

4.5 Interpretation of Schur Procedure

The theorem rests mainly on the concept of Schur complements and a lemma due to Schur. Let us write $\mathbf{R}_{0:N}$ as,

$$\mathbf{R}_{0:N} = \begin{bmatrix} R_{0,0} \\ R_{0,1} \\ \vdots \\ R_{0,N} \end{bmatrix} R_{0,0}^{-1}[R_{0,0} \ \cdots \ R_{0,N}] + \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \Delta & \\ 0 & & \end{bmatrix} \tag{88}$$

Then Δ is called the *Schur complement* of $R_{0,0}$ in $\mathbf{R}$. Then the following result is easy to see.

Lemma: $\mathbf{R}$ is positive definite if and only if $R_{o,o} > 0$ and the Schur complement of $R_{o,o}$ in $\mathbf{R}$ is positive definite.

The following reformulation of this fact will be useful. Let us define

$$\mathbf{S}(z,\omega) = \sum_{i,j=0}^{\infty} R_{i,j} z^i \omega^{*j} \tag{89}$$

If $\mathbf{R} = [R_{i,j}]$ is positive definite, this function will be said to be the *two-dimensional spectral function of the process*. (Caution: The spectral function $\mathbf{S}(z,\omega)$ must not be confused with the two dimensional Schur function $S(z,\omega)$ suggested in (72)). Then we can restate the lemma as follows:

Lemma: The function $\mathbf{S}(z,\omega)$ associated with $\mathbf{R}$ is a spectral function if and only if $\mathbf{S}(0,0) = R_{o,o} > 0$ and $\mathbf{S}_1(z,\omega)$ is also a spectral function, where

$$\mathbf{S}_1(z,\omega) = \frac{\mathbf{S}(z,\omega) - \mathbf{S}(z,0)\mathbf{S}^{-1}(0,0)\mathbf{S}(0,\omega)}{z\omega^*} . \tag{90}$$

The reader should check, by comparing coefficients of $z^i\omega^{*j}$ that $\mathbf{S}_1(z,\omega)$ indeed represents the Schur complement Δ in the transform domain.

The above lemma may be extended to the "ARMA case" readily, which we do here in order to bring out connections to the paper of Dewilde in this volume.

Lemma (Schur-ARMA): Let "α" be point in the domain of convergence of $\mathbf{S}(z,\omega)$. Then $\mathbf{S}(z,\omega)$ is a spectral function if and only if $\mathbf{S}(\alpha,\alpha) > 0$ and $\mathbf{S}_1(z,\omega)$ is also a spectral function, where

$$\mathbf{S}_1(z,\omega) = \frac{\mathbf{S}(z,\omega) - \mathbf{S}(z,\alpha)\mathbf{S}^{-1}(\alpha,\alpha)\mathbf{S}(\alpha,\omega)}{\left[\dfrac{z-\alpha}{1-\alpha^* z}\right]\left[\dfrac{\omega-\alpha}{1-\alpha^*\omega}\right]^*} \tag{91}$$

These lemmas are of little practical use as it is, because computation of $\mathbf{S}_1(z,\omega)$ from $\mathbf{S}(z,\omega)$ is not efficient. So, we now try to impose a certain structure on $\mathbf{S}(z,\omega)$, which will hopefully enable us to perform recursions on some simpler function than $\mathbf{S}(z,\omega)$.

a. Schur Invariance:

First, let us look for a $\mathbf{S}(z,\omega)$ that remains unaltered under Schur complementation, i.e., such that

$$\mathbf{S}_1(z,\omega) = \mathbf{S}(z,\omega) \tag{92}$$

Substituting this into (91) gives, after some algebra, the fact that

$$(1 - z\omega^*)\mathbf{S}(z,\omega) = G_a(z)G_a^*(\omega) \tag{93a}$$

where

$$G_a(z) = \frac{(1-a^*z)\mathbf{S}(z,a)\mathbf{S}^{-1/2}(a,a)}{(1-|a|^2)^{1/2}} \tag{93b}$$

Let us pause a little here to examine the AR-case. Then $a = 0$, and (93) becomes

$$(1 - z\omega^*)\mathbf{S}(z,\omega) = G(z)G^*(\omega) \tag{94}$$

where

$$G(z) = \mathbf{S}(z,0)\mathbf{S}^{-1/2}(0,0)$$

In the "time-domain", (94) translates to the relation

$$\mathbf{R} - Z\mathbf{R}Z^* = \mathbf{G}\mathbf{G}^* \tag{95}$$

Processes satisfying (94), i.e., such that

$$\mathbf{S}(z,\omega) = G(z)\frac{1}{1-z\omega^*}G^*(\omega) \tag{96}$$

have a simple interpretation. Consider "one-sided" white noise $\{u_k,\ k \geq 0\}$

such that $Eu_k u_l^T = I\delta_{kl}$. Its spectral function is

$$\mathbf{S}_u(z,\omega) = \sum_{i,j=0}^{\infty} \delta_{ij} z^i \omega^{*j} = \frac{1}{1 - z\omega^*} \tag{97}$$

Now the process in (96) can be interpreted as the nonstationary process obtained by putting "one-sided" white noise through a time-invariant filter with transfer function $G(z)$. These are perhaps the simplest kind of nonstationary processes, with displacement rank $\alpha = 1$.

Let us now return to our main problem, of imposing certain useful structure on the processes to be considered. Our earlier condition that $\mathbf{S}(z,\omega)$ remain unaltered under Schur complementation, which resulted in $\mathbf{S}(z,\omega)$ taking the form in (96), can be generalized as follows:

Let

$$\mathbf{S}(z,\omega) = \frac{G_a(z)\Sigma G_a^*(\omega)}{1 - z\omega^*} \tag{98}$$

where $G_a(z)$ will be required to be "proper" in the sense that

$$G_a(a) = \sigma_a[1 \ 0 \ \cdots \ 0] \tag{99a}$$

and

$$\sigma_a^2 = \mathbf{S}(a,a)(1 - |a|^2) \tag{99b}$$

The Schur complement is given as below for the ARMA-case:

$$\mathbf{S}_1(z,\omega) = \frac{\mathbf{S}(z,\omega) - \mathbf{S}(z,a)\mathbf{S}^{-1}(a,a)\mathbf{S}(a,\omega)}{\left(\frac{z-a}{1-a^*z}\right)\left(\frac{\omega - a}{1 - a^*\omega}\right)^*}$$

Substituting (98) into the expression (91) shows that $\mathbf{S}_1(z,\omega)$ must have the form

$$\mathbf{S}_1(z,\omega) = \frac{G_1(z)\Sigma G_1^*(\omega)}{1 - z\omega^*} \tag{100a}$$

where

$$G_1(z) = G_a(z)\begin{bmatrix} 1 & 0 \\ 0 & \left(\frac{z-a}{1-a^*z}\right)I \end{bmatrix}^{-1} \Theta^{-1}(K_1,J) \tag{100b}$$

where K_1 is such that

$$G_1(a_2) = \sigma_{a_2}[1 \quad 0 \quad \cdots \quad 0] \tag{100c}$$

Thus by considering processes whose spectral function can be written as in (100a) we are able to reduce recursions of $S(z,\sigma)$ to recursions on $G_a(z)$. The G_a-recursions are related through Θ-matrices. So, we can continue this above procedure by extracting Θ-sections at $a_2, a_3, \cdots$. Note that after one step, we have

$$G(z) = G_1(z)\begin{bmatrix} I & 0 \\ 0 & \left(\frac{z-a}{1-a^*z}\right)I_{a-1} \end{bmatrix} \Theta(K_1,J) \tag{101a}$$

$$= G_1(z) \cdot \Theta_1(z) \ . \tag{101b}$$

The process will be *finite order* if

$$G(z) \cdot \Theta_1^{-1}(z) \quad \cdots \quad \Theta_M^{-1}(z) = G_M(z)[1 \quad 0 \quad \cdots \quad 0] \quad \text{for} \quad M < \infty. \tag{102}$$

We shall make this assumption for convenience, although it is not essential. Thus,

$$G(z) = G_M(z)[1 \quad 0 \quad \cdots \quad 0]\Theta_M(z) \quad \cdots \quad \Theta_1(z) \tag{103}$$

Note that the product of Θ matrices is again a Σ-lossless matrix.

b. Congruence:

Here we will show that the Schur coefficiens K_i (and hence Θ - sections) do not uniquely determine a process. In fact, let $h(z)$ be analytic and invertible in some region containing the points $\alpha_1, \alpha_2, \ldots \alpha_N$. Then $G(z)$ and $h(z)G(z)$ have the same Schur coefficients. For,

$$h(z)G(z) = [h(z)\gamma(z) \quad h(z)\Gamma(z)]$$

and since we work with ratios, it is suggestive that the Schur coefficients will be the same. So, $G(z)$ and $h(z)G(z)$ are said to be *congruent*. The following lemma shows the relation between covariance matrices of congruent processes.

Lemma: If $G(z)$ is replaced by $h(z)G(z)$, i.e., $G(z) \to h(z)G(z)$, then the covariance matrix is transformed as:

$$\mathbf{R}_{o:N} \to L(h)\mathbf{R}_{o:N}L^*(h) \ , \qquad L(h) := \begin{bmatrix} h_o & & & 0 \\ h_1 & \cdot & & \\ \vdots & & \cdot & \\ h_N & \cdots & h_1 & h_0 \end{bmatrix} \tag{104}$$

Proof: In the time domain,

$$G(z) \to h(z)G(z) \ , \text{ implies that } G \to L(h)G \ .$$

Now from

$$\mathbf{R} - Z\mathbf{R}Z^* = G\Sigma G^*$$

we can write

$$L\mathbf{R}L^* - LZ\mathbf{R}Z^*L^* = LG\Sigma G^*L^*$$

Since $ZL = LZ$, we have

$$L\mathbf{R}L^* - Z(L\mathbf{R}L^*)Z^* = LG\Sigma G^*L^* \ .$$

which shows that, under congruence of the generators, we have congruence of the covariance matrices, i.e.,

$$\mathbf{R} \to LRL^* .$$

Since the K-coefficients define a covariance only up to congruence, we shall now consider a canonical representative.

c. Canonical Representative:

It can be shown that every generator $G(z)$ can be transformed by a congruence into a canonical form.

$$\bar{G}(z) := [1 \quad 0 \quad \cdots \quad 0]\Theta_M(z) \quad \cdots \quad \Theta_1(z) \tag{105}$$

Note that $\bar{G}(z)$ is completely determined by the Schur coefficients.

So far we have spoken only of tests for positivity and of the generalized Schur coefficients. Through the canonical generators we shall be able to relate these to the prediction or whitening filters that define the residuals $\{e_{m,t}\}$.

4.6 Innovations Filters and Generalized Orthogonal Polynomials

As we just stated there is a transformation relating $G(z)$ to $\bar{G}(z)$, which we can write as,

$$a(z)G(z) = \bar{G}(z) \tag{106a}$$

Let $G_{o:m}$ denote the first m rows of $G_{o:M}$. Then we can find a polynomial $a_m(z)$ of degree m such that

$$a_m(z) \cdot G(z) = \bar{G}_m(z) + O(z^{m+1})$$

We shall write this as

$$a_m(z)G(z) \overset{m}{=} \bar{G}_m(z) \tag{106b}$$

where $\overset{m}{=}$ denotes equality of the first m terms. Then after some calculation, for which we refer to the thesis of Lev-Ari (1982), we can claim the following.

56

THEOREM: The function $a_m(z)$ defines the innovations filter at time $t = m$.

This theorem does not indicate how to actually compute the $\{a_m(z)\}$. To do this conveniently, we introduce the concept of *admissibility*.

ADMISSIBILITY: A generator $G(z)$ is called λ-*admissible* if

$$G(z)\lambda = 1 \tag{107}$$

where λ is a vector of dimension α, i.e.,

$$\begin{bmatrix} g_0 \\ \cdot \\ \cdot \\ \cdot \\ g_N \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_\alpha \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

[Lerer and Tismenetsky (1981) introduced similar constraints to study generalized Bezoutians.] We know that

$$a_m(z)G(z) \overset{m}{=} \bar{G}_m(z)$$

Multiplying both sides by λ, and using (107) gives us $a_m(z)$ as

$$\begin{aligned} a_m(z) &= \bar{G}_m(z)\lambda \ . \\ &= [1 \quad 0 \quad \cdots \quad 0]\Theta_m(z) \quad \cdots \quad \Theta_1(z)\lambda \end{aligned} \tag{108}$$

We can represent this in block diagram form as shown in Figure 4. Each section in this cascade structure has the generalized lattice form shown in Figure 3.

557

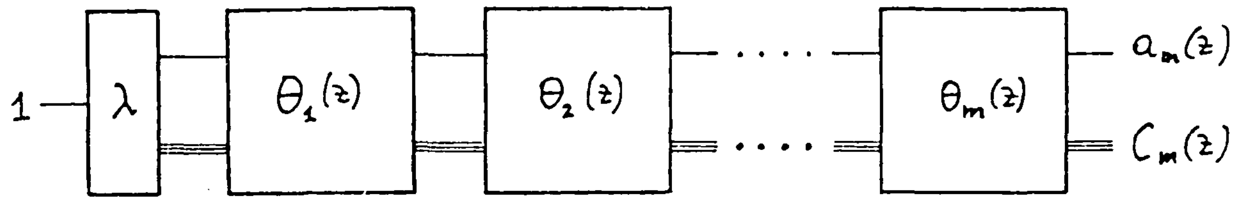


Figure 4. Generalized Lattice Filter with Time-Invariant Sections

Orthogonal Polynomials

The theorem that the $a_m(z)$ define the prediction-error filters means that they have certain orthogonality properties. In fact, they turn out to be natural extensions of the Szegö polynomials orthogonal on the unit circle (see, e.g., Geronimus (1961), Kailath, Vieira and Morf (1978)). If we denote the other outputs from the cascade $\Theta_m(z),\ldots,\Theta_1(z)$ by $C_m(z)$, an $(\alpha-1)$-dimensional vector, then we clearly have the recursive updating formula

$$\begin{bmatrix} a_m(z) \\ C_m(z) \end{bmatrix} = \Theta_m(z) \begin{bmatrix} a_{m-1}(z) \\ C_{m-1}(z) \end{bmatrix} \qquad (109)$$

which generalizes the well known Szego recursion formula. The formula (109) is (a slight generalization of) the generalized Levinson recursions obtained by Friedlander, Kailath, Morf, Ljung (1978), (1979). The present derivation is much more transparent, and in fact it was in large part the search for such a simple derivation that led to the studies presented here.

Christoffel-Darboux Formulas

With orthogonal polynomials, we always have the so-called Christoffel-Darboux formulas, which give expressions for $\mathbf{R}^{-1}$. In our case, we have the following generalized forms.

Let

$$K_m(z,\omega) := [1 \;\; z \;\; \cdots \;\; z^m]\mathbf{R}_{0:m}^{-1}[1 \;\; \omega \;\; \cdots \;\; \omega^m]^* \tag{110}$$

and also define the "reciprocal polynomial" by

$$\hat{a}_m(z) = z^m[a_m(1/z^*)]^* \tag{111}$$

Then we can establish the generalized Christoffel-Darboux formulas,

$$K_m(z,\omega) = \sum_{i=0}^{m} \hat{a}_i(z)\hat{a}_i^*(\omega) \tag{112}$$

$$= \hat{a}_{m+1}(z)\hat{a}_{m+1}^*(\omega) - \hat{C}_{m+1}(z)J\hat{C}_{m+1}^*(\omega) + \lambda^*\Sigma\lambda \tag{113}$$

The only surprising thing about this generalized formula is the presence of the term $\lambda^*\Sigma\lambda$. Some light will be shed on this in the next subsection.

4.7 More on Generators and Admissibility

Now we address ourselves to the question, can we always find admissible generators? The answer is essentially 'yes', which we shall elaborate now. First, we note the concept of equivalence and certain results due to Livsic (1979) and Potapov (1960). Equivalent generators $\{G_1,\Sigma_1\}$ and $\{G_2,\Sigma_2\}$ are such that

$$\mathbf{R} - Z\,\mathbf{R}\,Z^* = G_1\Sigma_1G_1^* = G_2\Sigma_2G_2^*$$

The following results are an adaptation of results proved by Livsic and Potapov.

1. All minimal generators are related by Σ-unitary matrices.

2. In general, all generators can be obtained from any minimal generator by some combination of certain elementary operations (called trivial lengthenings, neutral lengthenings, and Σ-rotations).

Now let us consider the problem of finding admissible generators. It turns out that we have three distinct cases, which in fact give a classification of covariances into three distinct types: Let us define the index of non-Toeplitzness of $\mathbf{R}$ as first introduced in Friedlander et al (1978) to be (see (114) below)

$$\delta = \text{rank}\,[R_{t,s} - R_{t-1,s-1}]_{t,s\geq 1}$$

It is easy to see that $(\alpha-\delta)$ can be 2, 1 or 0, where of course α is the displacement rank.

Theorem: Let $\mathbf{R}-Z\,\mathbf{R}\,Z^*=G\Sigma G^*$ where G is a minimal generator. Let

$$G(z) = \sum_i g_i z^i$$

Then for *every* minimal generator, if

(i) $\alpha-\delta=2$, there exists a λ such that $G(z)\lambda=1$ and $\lambda^*\Sigma\lambda=0$. An example is the Toeplitz case, where $\lambda=[1\ \ -1]$.

(ii) $\alpha-\delta=1$, there exists a λ such that $G(z)\lambda=1$, but $\lambda^*\Sigma\lambda\neq 0$

(iii) $\alpha-\delta=0$, there exists *no* λ such that $G(z)\lambda=1$. However, by increasing the dimension to $(\delta+2)$, we can always find a G that is admissible, with *any* desirable λ. An example is the pre-windowed least-squares problem of Sec. 3.4.

Remark: The class of covariances satisfying $\alpha-\delta=2$ deserves the name ***quasistationary*** or close to stationary. The simplest members in this class are, indeed, Toeplitz covariances, and for any other member the distance from stationarity is indicated by the index of non-Toeplitzness $\delta = \alpha - 2$. Notice also that the surprising term $\lambda^* \Sigma \lambda$ which appeared in (113) vanishes for covariances in this class.

Construction of an Admissible Generator

We shall give a method of constructing an admissible generator, though it may not necessarily be minimal. Let

$$\mathbf{R} - Z\,\mathbf{R}\,Z^* = \begin{bmatrix} R_{0,0} & R_{0,1} \; . \; . \; . & R_{0,N} \\ R_{1,0} & & \\ \cdot & & \\ \cdot & \tilde{\Delta} & \\ \cdot & & \\ R_{N,0} & & \end{bmatrix} \tag{114}$$

we write $\tilde{\Delta}$ as,

$$\tilde{\Delta} = D_{1:N} \Sigma_0 D_{1:N}^*$$

where Σ_0 has dimension $\delta \times \delta$. Now define,

$$G_{0:N} = \left[\begin{array}{c|ccc|c} \overline{R}_{0,0} & 0 & \cdots & 0 & 0 \\ \overline{R}_{1,0} & & & & -\overline{R}_{1,0} \\ \cdot & & D_{1:N} & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \overline{R}_{N,0} & & & & -\overline{R}_{N,0} \end{array}\right] \quad \text{where} \quad \overline{R}_{i,o} = R_{o,o}^{-1/2} R_{i,o}$$

and

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \Sigma_o & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Then, it is easily checked that

$$G_{o:N} \Sigma G_{o:N}^* = \mathbf{R} - Z\mathbf{R}Z^*$$

and that

$$G_{o:N} \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

thus making $G_{o,N}$ an admissible generator, with $\lambda^* = [1 \quad 0 \quad \cdots \quad 0 \quad -1]$.

Remark: Note that this construction matches the statement in the theorem when $\alpha - \delta = 2$ or $\alpha - \delta = 0$. It is however, nonminimal if $\alpha - \delta = 1$.

4.8 Concluding Remarks

We have given an outline of the general theory of lattice filters for nonstationary processes, emphasizing how the displacement structure of the covariance can be used to obtain different kinds of implementations. We find it fascinating that the displacement rank concept provides such technologically significant generalizations of the several classical results of Schur and Szego, obtained without any idea of practical applications.

Acknowledgments: It is a pleasure to thank A. Benveniste, P. Dewilde and Y. Landau for their care and dedication in organizing a fine workshop, providing this opportunity to present an outline of some very new research at Stanford. I am grateful to R. Prabhakhar of Delft for his notes of the oral presentation, which I have reworked here, and of course also to my several colleagues at Stanford, whose work over the last few years formed the basis for the oral presentation.

REFERENCES

A. Alam (1978), "Orthonormal Lattice Filters - a Multistage Multichannel Estimation Technique", *Geophysics*, vol. 43, pp. 1368-1383.

J. Burg (1967), "Maximum Entropy Spectral Analysis", presented at the 37th Ann. Int. Meeting Soc. of Explor. Geophys., Oklahoma City, OK.

J. Burg (1975), "Maximum Entropy Spectral Analysis", Ph.D. Dissertation, Stanford University, Stanford, CA.

J. F. Claerbout (1976), *Fundamentals of Geophysical Data Processing with Applications to Petroleum Prospecting*, McGraw-Hill, New York.

J. Delosme and M.Morf (1980), "A Tree Classification of Algorithms for Toeplitz and Related Equations", *Proc. 19th IEEE Conf. on D&C*, pp. 42-46, Albuquerque, New Mexico.

P. Dewilde (1969), "Cascade Synthesis of Passive Networks", Ph.D. Dissertation, Dept. of Electrical Eng., Stanford University, Stanford, CA.

P. Dewilde, A. C. Vieira and T. Kailath (1978), "On a Generalized Szegö-Levinson Realization Algorithm for Optimal Linear Predictors Based on a Network Synthesis Approach", *IEEE Trans. Circuits and Systems*, vol. CAS-25, no. 9, pp. 663-675.

A. Fettweis, H. Levin and A. Sedlmeyer (1974), "Wave Digital Lattice Filters", *Int. J. Circuit Thy. and Appl.*, vol. 2, pp. 203-211.

B. Friedlander, T. Kailath, M. Morf and L. Ljung (1978), "Extended Levinson and Chandrasekhar Equations for General Discrete-Time Linear Estimation Problems", *IEEE Trans. Autom. Contr.*, vol. AC-23, no. 4, pp. 653-659.

B. Friedlander, M. Morf, T. Kailath and L. Ljung (1979), "New Inversion Formulas Matrices Classified in Terms of Their Distance from Toeplitz Matrices", *Linear Algebra and Its Appls.*, vol. 27, pp. 31-60.

Y. Geronimus (1961), *Orthogonal Polynomials*, Consultant Bureau, New York (Russian original, 1956).

W. Hodgkiss and Presley (1981), "Adaptive Tracking of Multiple Sinusoids Whose Power Levels are Widely Separated", *IEEE Trans. Acous. Speech and Signal*

Proc., vol. ASSP-29, pp. 710-721.

F. Itakura and S. Saito (1971), "Digital Filtering Techniques for Speech Analysis and Synthesis", *Proc. 7th Int. Conf. Acoust.*, Paper 25-C, pp. 261-264, Budapest.

T. Kailath (1973), "Some New Algorithms for Recursive Estimation in Constant Linear Systems", *IEEE Trans. on Inform. Thy.*, vol. IT-19, no. 6, pp. 750-760.

T. Kailath, S. Kung and M. Morf (1979), "Displacement Ranks of Matrices and Linear Equations", *J. Math. Anal. Appl.*, vol. 68, no. 2, pp. 395-407.

T. Kailath, L. Ljung and M. Morf (1978), "Generalized Krein-Levinson Equations for Efficient Calculation of Fredholm Resolvents of Nondisplacement Kernels", *Topics in Functional Analysis*, Essays in honor of M. G. Krein, Advances in Math. Supplementary Studies, vol. 3, Academic Press, New York.

T. Kailath, G. S. Sidhu and M. Morf (1973), "Some New Algorithms for Recursive Estimation in Constant Linear Discrete-Time Systems", *Proc. Seventh Princeton Conf. on Inform. Sci. & Systs.*, pp. 344-352, Princeton, N.J.

T. Kailath, A. Vieira and M. Morf (1978), "Inverses of Toeplitz Operators, Innovations, and Orthogonal Polynomials", *SIAM Rev.*, vol. 20, no. 1, pp. 106-119.

D. Lee (1980), "Canonical Ladder Form Realizations and Fast Estimation Algorithms", Ph.D. Dissertation, Dept. of Electrical Eng., Stanford University, Stanford, CA.

D. Lee, M. Morf and B. Friedlander (1981), "Recursive Least Squares Ladder Estimation Algorithms", *IEEE Trans. Circuits and Systems*, vol. CAS-28, pp. 467-481.

L. Lerer and M. Tismenetsky (1981), *The Bezoutian and the Eigenvalue Separation Problem for Matrix Polynomials*, Technion Report, Haifa, Israel.

H. Lev-Ari (1982), "Parametrization and Modeling of Nonstationary Processes", Ph.D. Dissertation, Dept. of Electrical Eng., Stanford University, Stanford, CA.

M. Livsic and A. A. Yantsevich (1979), *Operator Colligations in Hilbert Spaces*, J. Wiley & Sons, New York.

J. Makhoul (1977), "Stable and Efficient Lattice Methods for Linear Prediction", *IEEE Trans. Acoust., Speech and Signal Proc.*, vol. ASSP-25, pp. 423-428.

J. D. Markel and A. H. Gray, Jr. (1976), *Linear Prediction of Speech*, Springer-Verlag, New York.

M. Morf (1974), "Fast Algorithms for Multivariable Systems", Ph.D. Dissertation, Dept. of Electrical Eng., Stanford University, Stanford, CA.

M. Morf and D. T. Lee (1978), "Fast Algorithms for Speech Modeling", Tech. Rept. No. M303-1, Information Systems Lab., Stanford University, Stanford, CA.

M.Morf and D. T. Lee (1979), "Recursive Least Squares Ladder Forms for Fast Parameter Tracking", *Proc. 1978 IEEE Conf. D&C*, pp. 1328-1367, San [illegible] CA.

M. Morf, D. T. Lee, J. R. Nickolls and A. Vieira (1977), "A Classification of Algorithms for ARMA Models and Ladder Realizations", *Proc. 1977 IEEE Conf. on Acoust., Speech and Signal Processing*. pp. 13-19, Hartford, CT.

M. Morf, A. Vieira and D. Lee (1977), "Ladder Forms for Identification and Speech Processing", *Proc. 1977 IEEE Conf. D&C, pp. 1074-1078, New Orleans, LA.*

B. Porat, B. Friedlander and Morf (1982), "Square-Root Normalized Covariance Algorithms", *IEEE Trans. Autom. Contr.*, to appear Aug. 1982.

Yu. Potapov (1960), "The Multiplicative Structure of J-Contractive Matrix Functions", *American Math. Soc. Transl.* Ser. 2, 15, pp. 131-244.

E. A. Robinson (1967), *Multichannel Time-Series Analysis with Digital Computer Programs*, Holden-Day, San Francisco, CA.

E. Satorius and J. Pack (1981), "Application of Least Squares Lattice Algorithms to Adaptive Equalization", *IEEE Trans. Comm.*, vol. COM-29, pp.136-142.

I. Schur (1917), "Uber Potenzreihen, die in Innern des Einheitskreises beschrankt sind", *J. fur Reine und Angew. Math.*, vol. 147, pp. 205-232.

G. Sidhu and T. Kailath, "A Shift-Invariance Approach to Chandrasekhar, Cholesky and Levinson-Type Algorihms", *Proc. 1975 Johns Hopkins Conf. on Inform. Sci. and Systs.*, pp. 324-327, Balimore, MD.

G. Szegö (1939), *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Pub., vol. 23, Providence, R.I.

G. W. Stewart (1973), *Introduction to Matrix Computations*, Academic Press, New York.

A. C. Vieira (1977), "Matrix Orthogonal Polynomials with Applications to Autoregressive Modeling and Ladder Forms", Ph.D. Dissertation, Dept. of Electrical Eng., Stanford University, Stanford, CA.

H. Wakita (1973), "Direct Estimation of the Vocal Tract Shape by Inverse Filtering of Acoustic Speech Waveforms", *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 417-427.

G. U. Yule (1907), "On the Theory of Correlation for any Number of Variables, Treated by a New System of Notation", *Proc. Roy. Soc.*, vol. 79A, pp. 182-193.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 82-0481	2. GOVT ACCESSION NO. AD-A116 243	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TIME-VARIANT AND TIME-INVARIANT LATTICE FILTERS FOR NONSTATIONARY PROCESSES		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Thomas Kailath		8. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0058
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Electrical Engineering Stanford University Stanford CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A6
11. CONTROLLING OFFICE NAME AND ADDRESS Mathematical & Information Sciences Directorate Air Force Office of Scientific Research Bolling AFB DC 20332		12. REPORT DATE MAY 82
		13. NUMBER OF PAGES 48
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Time-variant lattice filters; time-invariant lattice filters; nonstationary processes; whitening and modeling filters; Cholesky decompositions.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The structure of second-order processes is exposed by specification of whitening filters and modeling filters, or equivalently by Cholesky decompositions of the covariance matrix and its inverse. We shall show that these filters can be obtained as a cascade of lattice sections, each specified by a single so-called reflection coefficient parameter. For stationary processes, the reflection coefficient will be time-invariant. For nonstationary processes we can use the displacement rank concept either to find a simple time-update formula for the reflection coefficients or to replace them by a time-invariant vector reflection (CONTINUED)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ITEM #20, CONTINUED: coefficient of size governed by the displacement rank of processes.

These results are obtained in a quite direct way by using a geometric (Hilbert-space) formulation of the problem, combined with old results of Yule (1907) on update formulas for partial correlation coefficients and of Schur (1917) on the classical moment problem.

DATE
LMED

7-8